Does Public Opinion Affect Political Speech?

Hager, Anselm; Hilbig, Hanno

Hager, Anselm; Hilbig, Hanno

Does Public Opinion Affect Political Speech?

American Journal of Political Science

Provided in Cooperation with:
WZB Berlin Social Science Center

Suggested Citation: Hager, Anselm; Hilbig, Hanno (2020) : Does Public Opinion Affect Political Speech?, American Journal of Political Science, ISSN 1540-5907, Wiley, Oxford, Iss. Early View Articles,
http://dx.doi.org/10.1111/ajps.12516

This Version is available at:
http://hdl.handle.net/10419/218852

AJPS AMERICAN JOURNAL of POLITICAL SCIENCE

Does Public Opinion Affect Political Speech?

Anselm Hager Humboldt-Universität zu Berlin
Hanno Hilbig Harvard University

Abstract: *Does public opinion affect political speech? Of particular interest is whether public opinion affects (i) what topics politicians address and (ii) what positions they endorse. We present evidence from Germany where the government was recently forced to declassify its public opinion research, allowing us to link the content of the research to subsequent speeches. Our causal identification strategy exploits the exogenous timing of the research's dissemination to cabinet members within a window of a few days. We find that exposure to public opinion research leads politicians to markedly change their speech. First, we show that linguistic similarity between political speech and public opinion research increases significantly after reports are passed on to the cabinet, suggesting that politicians change the topics they address. Second, we demonstrate that exposure to public opinion research alters politicians' substantive positions in the direction of majority opinion.*

Verification Materials: The data and materials required to verify the computational reproducibility of the results, procedures, and analyses in this article are available on the *American Journal of Political Science* Dataverse within the Harvard Dataverse Network, at: https://doi.org/10.7910/DVN/SLXRVJ.

Does public opinion affect political speech? This important question of political science has received little empirical scrutiny. The handful of published studies have yielded conflicting findings. One influential study by Jacobs and Shapiro (2000) finds little evidence that public opinion meaningfully affected the rhetoric of President Bill Clinton. By contrast, Rottinghaus (2008) uses White House archival data and demonstrates that public opinion places significant constraints on presidential framing. Evidence that reliably establishes a causal connection between public opinion and political speech, however, is scarce.

What explains the lack of evidence? First, elected officials typically mask their means of gauging public opinion and their exposure thereto. Second, the dissemination of public opinion to politicians must be exogenous if one wants to establish a causal relationship. Third, when analyzing political speech, researchers face nontrivial measurement challenges. Any analysis into the interplay of public opinion and political speech must differentiate whether it explores i) the topics politicians address or ii) the substantive positions they endorse.

The present article circumvents these challenges by using an unusual source of information: classified governmental public opinion research. In 2013, a politician of the German Green Party sued the German government in order to gain access to its classified public opinion research. The German Federal Press Office (*Bundespresseamt*, henceforth BPA) subsequently and reluctantly granted the politician access to all public opinion research conducted between 2009 and 2013. Overall, the politician, together with a team of journalists, hand-copied more than 10,000 pages of classified public opinion research.

These research reports offer a unique lens into elected officials' exposure to public opinion. They are addressed directly to Germany's chancellor, Angela Merkel, and are disseminated to all cabinet members. The purpose of the reports is to give political elites a succinct overview of public opinion. The research is conducted by Germany's major public opinion research firms. Most

Anselm Hager is Assistant Professor of International Politics, Department of Social Sciences, Humboldt-Universität zu Berlin, Universitätsstrasse 3b, 10117, Berlin, Germany (anselm.hager@hu-berlin.de). Hanno Hilbig is PhD Candidate, Department of Government, Harvard University, 1737 Cambridge Street, Cambridge, MA 02138 (hhilbig@g.harvard.edu).

The authors are grateful to Sven Becker, Christian Breunig, Jacob Brown, Andy Eggers, Niklas Flamang, Florian Foos, Michael Gill, Robert Goodwin, Donald Green, Johannes Hemker, Macartan Humphreys, David Ifkovits, Yegor Lazarev, Gwyneth McClendon, Tamar Mitts, Giuseppe Piertrantuono, Sven-Oliver Proksch, Christian Rauh, Nils Redeker, Arthur Spirling, Dustin Tingley, and seminar participants at Harvard, Oxford, Konstanz, the Annual Meeting of the American Political Science Association 2017, and the Social Science Center Berlin for help and feedback during different stages of the project.

American Journal of Political Science, Vol. 00, No. 00, xxxx 2020, Pp. 1–17

DOI: 10.1111/ajps.12516

reports include up to 60 pages of qualitative and quantitative insights. It is here that cabinet members find a rich repertoire to attune, or not, their speech to public opinion.

How does exposure to public opinion affect politicians' speech? We hypothesize that exposure to public opinion can lead to two different reactions. First, it may lead politicians to change what topics they address (agenda setting; John and Jennings 2010; Mortensen et al. 2011). The logic is that the reports make specific topics cognitively salient and also signal what topics are of relevance to the electorate. Second, exposure to public opinion may lead politicians to adjust their substantive positions to the median voter (Eggers and Spirling 2014). After all, the research reports clearly communicate what the population thinks.

To explore these hypotheses, we take advantage of the fact that the public opinion reports indicate the date on which the BPA sent them to the cabinet. And, there are good reasons to believe that the precise timing is exogenous (within a window of a few days). The reports undergo a tedious tendering process, and we show that the dissemination dates do not cluster around salient events such as elections or parliamentary sessions. The dissemination timing is also orthogonal to the media salience of a given report's topic as well as to citizen satisfaction with the government.

Based on the plausibly exogenous timing of the reports' dissemination to cabinet members, our empirical strategy is to compare politicians' speech to the public opinion reports right before and right after a report was issued. We measure political speech using all published government speech documents from 2005 to 2016 ($>$20,000). To capture agenda setting, we measure linguistic (cosine) similarity between the reports and political speech. To explore substantive responsiveness, we measure agreement between the reports and political speech by hand-coding a random subset of 2,000 speech–report pairs.

Using a regression discontinuity design, we yield two key pieces of evidence. First, we find that linguistic similarity between politicians' speech and the public opinion reports increases by 0.014 points (S.E. = 0.007) on a 0-1 scale after reports are passed on to the cabinet. We interpret this finding to mean that exposure to public opinion changes what topics politicians address. Second, we find that substantive agreement increases by 0.2 points (S.E. = 0.07) on a 7-point scale. The finding thus implies that exposure to public opinion also leads politicians to become more responsive to the public's preferences.

Can our empirical strategy tell agenda setting and substantive responsiveness apart? One concern regarding the main findings is that the measures of agenda setting and responsiveness overlap. Specifically, one may object that if politicians adjust their substantive positions, this also manifests itself in increased linguistic similarity. Although this measurement concern does not compromise the causal inferences we draw, it does highlight the difficulty of adjudicating between agenda setting and substantive repositioning. To parse the two strategies apart, we present two pieces of evidence. First, we show that both measures are not correlated. Second, we demonstrate that effect heterogeneity by topic salience is different across the two measures. When politicians address salient topics, we only see increases in similarity, while substantive agreement does not change. By contrast, when politicians address nonsalient topics, we see increased substantive agreement, but no changes in cosine similarity. The finding helps clarify the conditions under which politicians change their agendas and when they decide to take different substantive positions: When topics are salient, substantive realignment may be too risky, which makes changing agendas more attractive. The reverse holds for nonsalient topics. Here, changing agendas carries little weight, but substantive repositioning may win over voters without risking politicians' reputation.

Our results contribute to three related debates in political science. First, we add to a growing literature analyzing political speech (Binzer Hobolt and Klemmensen 2008). Scholars have long argued that politicians' speech, particularly that of U.S. presidents, influences public opinion (Cohen 1999). We complement this literature by showcasing that causality also runs in the opposite direction. Though others have explored this relationship (Wood and Lee 2009; Rottinghaus 2006), we are the first to provide causal estimates.

Second, the evidence adds to a debate about whether executive leaders follow centrist or partisan considerations. By highlighting that German politicians adjust their speech to public opinion (i.e., the median voter), we support a centrist reading of executive political behavior (e.g., Canes-Wrone 2006). By the same token, the evidence rejects a partisan reading, whereby executives cater to their own political clientele (Wood 2009).

Last, we contribute to a growing literature that applies text-analytical methods to political science (Slapin and Proksch 2008; Lowe et al. 2011; Grimmer and Stewart 2013; Roberts et al. 2014; Lucas et al. 2015). Recent empirical studies have assessed, inter alia, bureaucratic transcripts (Egesdal, Gill, and Rotemberg n.d.) and diplomatic cables (Gill and Spirling 2015). Analyses regarding internal government reports, however, are few and descriptive in nature. We bring large-N causal evidence to this literature, showing that public opinion affects politicians' speech agenda and substantive positions.

Theoretical Background

Against the backdrop of a large literature underlining the rhetorical craftsmanship of elected officials (see Chong and Druckman 2007), little is known about the ways in which public opinion affects political speech. Vaughn and Villalobos, for instance, lament that social scientists have "ignored the determinants of what the president actually says" (2006, 681). Shapiro and Jacobs second that "we are only now beginning to learn about the relationship between presidents and the polling and public opinion analysis that has gone on in presidential administrations since the 1960s" (2001, 151). What is more, the few empirical accounts that assess whether elected officials adjust their speech to public opinion are inconclusive.

On the one hand, scholars have presented evidence that questions whether public opinion has any influence on political speech. In one historical study, Hall (2002) uses interviews and archival data to assess whether President George W. Bush used public opinion research to forge his rhetoric. The author finds that Bush had a distinct distaste for polls, labeling them "phony and artificial" (Hall 2002, 531). Similarly, Jacobs and Shapiro (2000) analyze President Clinton's failed health care reform campaign and Newt Gingrich's "Contract with America" and find little evidence for rhetorical pandering. One rare large-N study is provided by Wood and Lee (2009) who develop a measure of presidential liberalism from 1945 to 2005 and find presidents rather unresponsive to mass political preferences.

On the other hand, a different set of scholars has presented evidence demonstrating that public opinion does affect political speech. They point out the skillful employment of public opinion research in several U.S. administrations. Rottinghaus (2008), for instance, uses White House archival data, including internal polling reports, and finds that public opinion placed constraints on presidential framing of foreign policy. Relatedly, Geer (1996) finds that the professionalization of White House public opinion polling led presidents to adjust their speech to public preferences. In one systematic empirical study, Rottinghaus (2006) matches presidential statements spanning nine administrations to public opinion polling and demonstrates robust congruence.

A key limitation of the existing empirical studies is that they merely correlate politicians' speech with public opinion. As such, they fall short of providing causal estimates and simply reiterate the well-established finding that public opinion correlates with political speech. The prime reason for this shortcoming is that it is difficult to make a convincing case that elected officials' exposure to public opinion is exogenous. In addition, we lack empirical evidence from outside the United States or even the Oval Office.

At a more fundamental level, it is also unclear *how* public opinion affects politicians' speech. Providing elected officials with new information about public opinion may lead to two distinct reactions. First, obtaining up-to-date public opinion may inspire politicians to put a given topic on their agenda. Second, information on public opinion may lead politicians to substantively align their speech with the median voter. We discuss both pathways in turn.

Agenda Setting

At a basic level, exposure to public opinion may lead elected officials to change their agenda. We define such agenda setting as "the priority given to an issue" (Chong and Druckman 2007, 112). Put more simply, agenda setting captures whether an elected official talks about a specific issue or not. Why may exposure to public opinion affect elected officials' speech agenda? Two reasons are noteworthy.

First, the provision of public opinion may signal to politicians that a given topic is of relevance to citizens. Public opinion research—including the reports commissioned by the BPA—tends to focus on important issues. Public opinion is also known to be less volatile for salient issues (Weaver 1991). Thus, if a cabinet member is exposed to public opinion, she may interpret this as evidence that voters care about the topic, leading her to put the topic on her agenda.

Second, the provision of public opinion on a given topic makes the topic cognitively salient to the politician. And, like ordinary citizens, politicians try to "minimize the cognitive burdens for forming judgments by drawing on those considerations that are most accessible" (Koch 1998, 211). Thus, if a cabinet member reads a report on tax reform, she may draw on this very topic in her next speech or newspaper interview because the report made the topic cognitively salient.

H1 (*Agenda Setting*): Exposure to public opinion on a given topic leads politicians to put the topic on their agenda.

Responsiveness

Besides affecting politicians' speech agenda, exposure to public opinion may also lead politicians to change how they substantively address a given topic. More specifically, exposure to public opinion may spark (rhetorical)

responsiveness—a core tenet of representative democracy. We define responsiveness as "politicians follow[ing] preferences as they change" (Wlezien 2004, 2). Why may exposure to public opinion lead elected officials to endorse substantive positions that accord with the median voter?

First, public opinion research parses out what the majority thinks. Following Downs's median voter theorem, one would therefore expect that politicians craft their speech so as to cater to the majority (Wood and Lee 2009). While the Downsian assumptions do not hold in Germany's multiparty system, the main party studied here, Merkel's *Christian Democratic Union*, does consider itself centrist.

Second and related, the provision of public opinion research may prime elected officials to focus their attention toward the median voter. Traditional party politics leads politicians to think about their core clientele. Public opinion research, by contrast, analyzes the entire population. Besides providing quantitative information on majority opinion, this may thus lead politicians to adopt a more comprehensive view of representation.

H2 (*Responsiveness*): Exposure to public opinion on a given topic leads politicians to endorse majority opinion on the topic.

Data

Public Opinion Research Reports

We use an unusual source of evidence to study whether public opinion affects political speech. In September 2012, Malte Spitz, a member of the German Green Party, filed an official inquiry with the BPA. He demanded access to two research reports the BPA had commissioned on behalf of the German government. The office denied the request. Malte Spitz then sued the government under the German Freedom of Information Act. He was subsequently granted access to the reports but was only allowed to make hand copies (Becker and Horning 2014).

The party member, together with two *Der Spiegel* journalists, scanned 125 reports, which comprise over 10,000 pages. An exemplary report page is provided in Figure 1, showcasing their highly detailed nature. The research reports are all addressed directly to Chancellor Merkel ("Dear Mrs. Chancellor"), followed by a two-page summary written and signed by the head of the BPA. The purpose of these summaries is to condense the findings so as to give the chancellor and other cabinet members a succinct overview of public opinion. Most reports include up to 60 pages of qualitative and quantitative insights. It is here that cabinet members and speechwriters find a rich repertoire to attune speeches to public opinion.

The reports were commissioned by the BPA to leading German public opinion firms before being passed on to the German cabinet. They span the entire legislative term from 2009 to 2013. Importantly, the BPA is an independent German bureaucracy. Contracts with survey firms undergo a tedious tendering process (more

FIGURE 1 Exemplary Public Opinion Report

AMBIVALENTE BEWERTUNG DER STEUERENTLASTUNGEN

Seit Wochen diskutieren die Koalitionspartner wie die Öffentlichkeit kontrovers über die beschlossenen wie auch die noch geplanten Steuerentlastungen. Angesichts der Entwicklung der Staatsfinanzen werden verstärkt Bedenken geäußert, ob sich die in den Koalitionsvereinbarungen festgeschriebenen Entlastungspläne realisieren lassen. Diese Debatte geht an der Bevölkerung nicht spurlos vorüber. Während 2008 noch zwei Drittel der gesamten Bevölkerung überzeugt waren, dass es ohne weiteres möglich ist, die Steuern- und Abgabenlast deutlich zu senken, vertreten mittlerweile nur noch 50 Prozent diese Auffassung. In den neuen Bundesländern hat sich der Anteil der Bevölkerung, der von erheblichen Entlastungsspielräumen ausgeht, sogar von 74 auf 52 Prozent zurückgebildet.

Wachsende Zweifel an Entlastungsspielräumen

Frage: "Glauben Sie, dass es möglich wäre, die Steuern und Abgaben für die Bürger deutlich zu senken, oder glauben Sie das nicht?"

	Gesamtdeutschland		Westdeutschland		Ostdeutschland	
	2008	2009	2008	2009	2008	2009
	%	%	%	%	%	%
Wäre möglich	67	50	65	49	74	52
Glaube das nicht	24	38	25	38	19	37
Unentschieden, keine Angabe	9	12	10	13	7	11
	100	100	100	100	100	100

Basis: Bundesrepublik Deutschland, Bevölkerung ab 16 Jahre
Quelle: Allensbacher Archiv, IfD-Umfragen 10022 und 10048 © IfD-Allensbach

Note: The figure displays page 1 of an exemplary research report entitled "Population's View on Tax Burden and Taxation" (*Steuerbelastung und Steuersystem aus Sicht der Bevölkerung*). The English translation is as follows: "For weeks, the coalition partners and the public have been debating controversially over the implemented and planned tax reliefs. In light of the evolution of public finances, concerns have been raised as to whether the austerity plans set out in the coalition agreements can be achieved. This debate has left a mark on the population as well. While in 2008 two-thirds of the entire population were convinced that it would be possible to reduce the burden of taxes and levies, today only 50 percent still hold this view. In the Eastern states, the proportion of the population, which expects considerable relief margins, has even declined from 74 to 52 percent."

below). In the period of study, the BPA commissioned an average of 150 surveys per year. In so doing, the BPA spent an average of two million euros per year on its public opinion research (Becker and Hornig 2014). Within the BPA and during the period of study, public opinion research was overseen by the head of Unit 204 ("public opinion research and evaluation"), Ute Molitor, and the presentation of the findings to the cabinet was overseen by the head of the BPA, Steffen Seibert. Importantly, the Merkel government—unlike the chancellors Kohl and Schröder—specifically decided to commission reports from all major public opinion firms in order to receive a more objective overview of public opinion. That said, Merkel and the cabinet do not directly communicate with the research firms (a detailed discussion of exogeneity is provided later; Becker and Hornig 2014). The reports thus represent a novel lens through which one can assess whether exposure to public opinion affects political speech.

To validate that the public opinion research reports constitute accurate measures of public opinion, we gained access to Germany's most prominent and longest-running opinion poll: the Politbarometer. The Politbarometer is commissioned by Germany's public television network (ZDF) and was started in 1977. It consists of regular (about 1.5 per month) surveys that serve to "poll the opinions and attitudes of eligible Germans with regard to current events and issues." During the second Merkel term, the ZDF commissioned 76 surveys with a total of 132,321 respondents. Based on these data, we make two points in Appendix 9 in the supporting information (SI). First, we show that 75% of the BPA research reports tackle topics that voters, at the time of the write-up, consider to be pressing issues. Second, we show that the relative frequency of the topics covered in the Politbarometer is highly comparable to the relative frequency of the topics addressed in the government research reports.

To convert the public opinion reports into text data, we followed standard operating procedures in the text analysis literature. Using the R package `tm`, we first removed numbers, punctuation, and white spaces from the reports. We then stemmed all reports by removing stop words, prefixes, and suffixes. On average, there were 1,683 words per report after stemming.

The descriptive statistics of the opinion reports are provided in the left column of Table 1. We asked two independent coders to assign all reports to political topics—a straightforward task given that most reports already include topic headlines. The coders were asked to categorize reports according to the 15 topics used by the German government (*Topic (original)*; more below). To streamline the analysis, we collapsed the 15 topics into seven categories (*Topic (aggregated)*) that map more intuitively onto German public opinion (details on the aggregation rule are provided in SI Appendix 1).

Table 1 also reports the year and month of the reports' dissemination to cabinet members. Both variables are fairly uniformly distributed. Finally, the table reports which of the seven leading German public opinion firms commissioned the report (*Opinion firm*). There are three key players—Allensbach, Dimap, and Emnid—who wrote 72% of all reports.

Political Speech

We measure political speech using all speech documents published by the German government from 2005 to 2016—with Merkel in power during the entire period. These documents include speeches, press releases, and articles written by cabinet members. We scraped the documents from the website of the German government (www.bundesregierung.de). The resulting data set covers all published speech documents by cabinet members. We applied the same transformation to the text data, that is, removing superfluous characters and stemming.

The descriptive statistics of the speech data are given in the right column of Table 1. Again, we classify the topic of the speech (*Topic (original)*). In particular, the German government assigned 47% of all documents tags according to the content of the speech. As before, we collapsed these 15 topics into seven categories. These include foreign policy (28%), culture (19%), economic policy (15%), social policy (13%), environmental policy (9%), and interior and education (both 8%). The rest of the documents are published without an explicit topic assignment. To avoid discarding valuable information, we use a machine learning approach to automatically classify the untagged documents. Using a support vector machine that takes the tf-idf document-term matrix of the speech document corpus as the input, we achieve an out-of-sample prediction accuracy of 89% (for details, see SI Appendix 2). Table 1 also reports the month and year of the publication. As can be seen, publications increased notably in 2011. The month of the publication, on the other hand, is fairly homogeneous. Finally, Table 1 also indicates the speech documents' type, separating articles (59%/46%), press releases (36%/27%), and speeches (5%/27%).

Empirical Strategy

Does public opinion affect political speech? If so, does public opinion merely alter politicians' speech agenda or

Table 1 Descriptive Statistics

	Opinion Reports		Government Speech (Manually Classified)		Government Speech (Machine Classified)	
	Absolute	*%*	*Absolute*	*%*	*Absolute*	*%*
Total	**125**	**100**	**9,831**	**100**	**11,973**	**100**
Topic (original)						
Labor/Welfare	32	20	509	5.177	—	—
Foreign policy	23	14.375	2,168	22.053	—	—
Education/Research	3	1.875	770	7.832	—	—
Agriculture	2	1.25	386	3.926	—	—
Families	11	6.875	437	4.445	—	—
Finances	25	15.625	458	4.659	—	—
Health	4	2.5	345	3.509	—	—
Interior	5	3.125	555	5.645	—	—
Justice	3	1.875	212	2.156	—	—
Culture	6	3.75	1860	18.92	—	—
Environmental policy	7	4.375	537	5.462	—	—
Infrastructure	0	0.0	220	2.238	—	—
Defense	9	5.625	273	2.777	—	—
Economic policy/Energy	30	18.75	785	7.985	—	—
Economic development	0	0.0	316	3.214	—	—
Topic (aggregated)						
Culture	6	3.75	1,860	18.92	863	7.208
Economic policy	55	34.375	1,463	14.881	1,684	14.065
Education	3	1.875	770	7.832	631	5.27
Environmental policy	9	5.625	923	9.389	860	7.183
Foreign policy	32	20.0	2757	28.044	5406	45.152
Interior	8	5.0	767	7.802	1367	11.417
Social policy	47	29.375	1,291	13.132	1,162	9.705
Month						
January	14	11.2	740	7.527	983	8.21
February	6	4.8	774	7.873	904	7.55
March	11	8.8	925	9.409	838	6.999
April	10	8.0	797	8.107	1,027	8.578
May	9	7.2	791	8.046	1,129	9.43
June	10	8.0	949	9.653	1,213	10.131
July	14	11.2	832	8.463	831	6.941
August	9	7.2	830	8.443	636	5.312
September	12	9.6	907	9.226	1,082	9.037
October	7	5.6	794	8.076	1,063	8.878
November	9	7.2	884	8.992	1,319	11.016
December	14	11.2	608	6.185	948	7.918

(Continued)

TABLE 1 Continued

	Opinion Reports		Government Speech (Manually Classified)		Government Speech (Machine Classified)	
	Absolute	*%*	*Absolute*	*%*	*Absolute*	*%*
Total	**125**	**100**	**9,831**	**100**	**11,973**	**100**
Year						
2005	0	0.0	8	0.081	0	0
2007	0	0.0	15	0.153	0	0
2008	0	0.0	58	0.59	0	0
2009	22	17.6	38	0.387	828	6.916
2010	29	23.2	221	2.248	2,358	19.694
2011	26	20.8	947	9.633	1,573	13.138
2012	25	20.0	1,535	15.614	742	6.197
2013	23	18.4	1627	16.55	722	6.03
2014	0	0	1,898	19.306	900	7.517
2015	0	0	2,017	20.517	1,017	8.494
2016	0	0	1,467	14.922	1,576	13.163
2017	0	0	0	0	2,257	18.851
Opinion firm						
Allensbach	34	27.2	—	—	—	—
Dimap	26	20.8	—	—	—	—
Emnid	30	24.0	—	—	—	—
FG	12	9.6	—	—	—	—
GMS	3	2.4	—	—	—	—
Polis	8	6.4	—	—	—	—
TNS	12	9.6	—	—	—	—
Medium						
Article	—	—	5,785	58.844	5,492	45.87
Press release	—	—	3510	35.703	3,219	26.885
Speech	—	—	536	5.452	3,262	27.245

Note: "Absolute" refers to absolute frequencies of categories. The label "%" refers to relative frequency in percent. For "Opinion firm," "Year," and "Month," the total number of observations is equal to the total number of opinion reports (i.e., N = 125). Each opinion report can have up to three topics. We display the summary statistics for the government speech documents separately for documents that were categorized by the government and for documents that were automatically classified (see SI Appendix 2). We only predict aggregated issue categories for untagged documents.

does it also affect the substantive positions they endorse? To address these questions, we use two distinct measurement strategies, which we introduce in turn. Thereafter, we discuss our estimation strategy.

Measurement

Agenda Setting. To assess whether the public opinion reports change what topics politicians address, we measure whether reports and politicians' speech become more linguistically similar after a report has been issued. Formally, this means creating a distance measure $Similarity_{i,j}$ for all public opinion reports i and all speech documents j. This gives $N^{\text{Report}} \times N^{\text{Speech}} = 125 \times (9,831 + 11,973) = 2,725,500$ observations. Our preferred measure of distance is cosine similarity—a simple and common measure of dissimilarity (e.g., Egesdal, Gill, and Rotemberg n.d.). It ranges from 0 (documents are entirely dissimilar) to 1 (documents are exactly the same).

Responsiveness. To assess whether the public opinion reports affect politicians' substantive positions, we measure substantive agreement between reports and politicians' speech. Our strategy, here, relies on human coding. First, we drew a random sample of 2,000 speech–report pairs (from the set of speech–report pairs that are on the same topic and were released within 120 days of each other; more below). Second, we asked trained research

assistants (who were blinded to the treatment status) to assign each of the pairs a score measuring substantive agreement. The score ranged from –3 (the politician strongly disagrees with public opinion displayed in the report) to 3 (the politician strongly agrees with public opinion). If politicians' speech bears no relation to public opinion, the research assistants assigned a score of 0 (more details are provided in SI Appendix 4). Importantly, Figure 5 shows that the hand-coded responsiveness measure is unrelated to cosine similarity. This builds trust that both measures tap distinct reactions (more below).

Estimation Strategy

In order to test whether exposure to public opinion affects political speech, we exploit the plausibly exogenous timing of the reports' dissemination to cabinet members. Specifically, we construct a regression discontinuity (RD) design that compares politicians' speech to the public opinion reports right before and after a report is disseminated. The empirical strategy falls into a broader class of studies that analyze attitudes just before and after exposure to new information (e.g., Franco, Grimmer, and Lim 2017).[1] We discuss exogeneity and the discontinuity specification in turn.

Exogeneity of Report Dissemination Timing. What determines when reports are sent to the German chancellery? As we outline in the following, there are nine reasons to believe that the precise timing—within a number of days—of the reports' dissemination to the cabinet is exogenous.

First, German law requires that contracts above 500 euros (as is the case for reports of this scope) undergo an official tendering process. Potential contracts must be widely advertised, and at least three competitive bids should be received. Since all seven major German public opinion firms are regularly contracted, competition is high. The length of the tendering process is thus tough to predict, making it difficult to strategically time the dissemination of findings to the cabinet.

Second, the commissioned reports comprise detailed academic research. They include large representative surveys and detailed focus group discussions. This adds an additional variance to the timing. To assess the influence of data collection effort on the timing of report dissemination in detail, SI Table A2 shows results from a regression of the number of days of data collection on the time it took the BPA to disseminate a report *after* data collection had already concluded. The table shows that the length of data collection alone explains about 44% of variation in the timing of report dissemination.

A third, related check regarding a possibly strategic timing of the dissemination of the reports to cabinet members concerns the timing of the survey evidence contained in the reports and the date of the dissemination. In SI Figure A6, we plot the number of days between data collection and dissemination to the chancellor's office. The majority of reports are passed on to the chancellor a few days after the last data were collected. This is further evidence that BPA bureaucrats do not strategically time the report dissemination.

Fourth, to quantitatively test for a possible influence of cabinet members on the dissemination timing of the reports, we examine whether reports on more salient topics tend to be disseminated earlier than reports on more peripheral topics. Perhaps cabinet members try to accelerate the release of reports that cover salient topics. We measure topic salience based on newspaper mentions of a given topic (for details, see SI Appendix 3). Reassuringly, SI Table A2 shows that there is no relationship between issue salience and dissemination timing.

Fifth, if the German chancellery does influence the timing of the report dissemination, one would expect them to do so ahead of important elections. To test this hypothesis, we assess whether reports are more likely to be issued right before state elections. SI Figure A.4 shows that this is not the case. If anything, we observe that reports are more likely to be issued *after* elections. The reason might be that research firms are particularly busy ahead of elections and therefore delay the release of the opinion reports to the BPA.

Sixth, in a similar vein the German chancellery could also try to accelerate the dissemination of reports before parliamentary sessions so as to improve their parliamentary speeches. One would therefore expect to see a greater number of reports just prior to days when the German parliament is in session. However, as SI Figure A.5 demonstrates, reports are not more likely to be released in the days preceding parliamentary sessions.

Seventh, if the dissemination date of reports is strategically chosen, this might show up in a nonuniform distribution of release dates across week days or months of the year. Some days and months (notably, Fridays as well as the months of July and August) tend to be less busy in the German political system. However, Table 1 demonstrates a highly regular timing across the year. This further underlines the reading that the BPA tries to provide

[1] Franco, Grimmer, and Lim (2017) study the effect of presidential appeals in the United States on public approval, constituents' policy preferences, constituents' perceived issue salience, and social media activity.

German political elites with a regular, objective view about public opinion.

Eight, we can assess whether the covariates of the speech documents are similar right before and after opinion research reports are issued. In SI Figure A.14, we show that opinion report releases induce no meaningful changes across a number of control variables, including a given speech's topic, length and medium.

Ninth, if the government strategically times the dissemination of the reports, it may do so during times when voters are dissatisfied with the government. Yet, in SI Figure A.13, we show that reports are *not* more likely to appear when voters express greater dissatisfaction with the government—as measured in the Politbarometer.

Regression Discontinuity Design. Having made the case that the precise dissemination timing of the public opinion reports to the cabinet is likely exogenous (within a few days), we leverage this fact by adopting a regression discontinuity design. Conceptually speaking, we restrict the analysis to a short period before and after the dissemination of reports to cabinet members. The assumption is that political speech that lies around the time of the dissemination has isomorphic potential outcomes. By comparing speech before and after the report dissemination, we can estimate the local average treatment effect (balance tests are presented in SI Figure A.14).

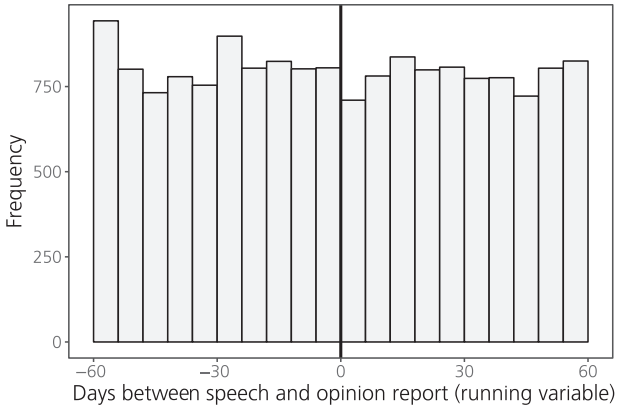
Our main empirical specification is as follows:

$$\begin{aligned} Y_{i,j} = \alpha + \tau \text{Exposure}_{i,j} + \beta_1 X_{i,j} \\ + \beta_2 X_{i,j} * \text{Exposure}_{i,j} + \zeta' \mathbf{Z}_{i,j} + \varepsilon_{i,j}. \quad (1) \end{aligned}$$

Here, $Y_{i,j}$ is speech document i's cosine similarity or substantive agreement, respectively, with report j. The running variable $X_{i,j}$ is the time in days between the release of the opinion report and the speech document, which is positive if the speech document is released after the report. $\mathbf{Z}_{i,j}$ is a vector of control variables listed in SI Table A6, and $\varepsilon_{i,j}$ is the error term. We cluster standard errors both by speech document and by opinion report.

Regression discontinuity designs require researchers to choose a bandwidth within which assignment to treatment is plausibly exogenous. Here, we rely on the bandwidth selection method proposed by Calonico, Cattaneo, and Titiunik (2014). To obtain the optimal bandwidth, we use a subsample of all report–speech pairs that are released at most 120 days apart and share the same topic. The resulting optimal bandwidth is 22 days.[2] Given that bandwidth calculations are not without criticism, we conduct sensitivity tests below.

FIGURE 2 Histogram of the Running Variable



Note: The figure plots a histogram of the running variable, that is, the number of days between speech document release and opinion report release (window of ±60 days).

Following Imbens and Lemieux (2008), we estimate the main equation using a local linear regression with the aforementioned optimal bandwidth of 22 days. More specifically, we fit local linear regressions using weights from a triangular kernel to estimate the treatment effect. Imbens and Lemieux (2008) show that the triangular kernel is optimal at the boundary.

A crucial assumption of RD designs is that the units of observation have no control over the assignment variable. In our case, this would be violated if politicians are able to time the release of a speech conditional on the publication date of an opinion report. To bolster this assumption, we follow McCrary (2008) who suggests that researchers estimate whether there is a discontinuity in the number of observations around the threshold. If the density of the assignment variable is not continuous, this may indicate that politicians adjust the timing of speeches to account for the release of opinion reports.

To test the "no jump"-assumption, Figure 2 presents a histogram of the running variable; that is, we plot the days between the speech document and the opinion report release for a window of 60 days around the release date. The figure shows that the number of observations before and after the cutoff is highly similar. There is thus no evidence of sorting. To formally test this assumption, we use a local polynomial density estimator proposed by Cattaneo, Jansson, and Ma (2019) and obtain an insignificant p-value of .67.

[2]We choose 120 days, rather than the maximum of 3,227 days, because the abovementioned algorithms otherwise choose bandwidths that are too large (400 days) to make credible inferences. We also confirm the main finding when using all available data—both in a simple OLS as well as when using a multilevel model (see SI Table A4).

Results

Agenda Setting

We begin by assessing whether exposure to the public opinion reports changes what topics politicians address. To do so, we assess whether linguistic similarity between public opinion reports and political speech increases after the dissemination of the reports. Table 2 demonstrates that the dissemination leads to a significant increase in linguistic similarity. Model 1 shows that the dissemination date increases cosine similarity between the reports and politicians' speeches by 0.014 (S.E. = 0.007) on a scale ranging from 0 to 1. In Model 2, we include all covariates provided in Table 1. The coefficient remains virtually unchanged. This builds trust that our setup is not compromised by unobserved confounding.

How large are the estimated effect sizes? As Table 2 reports, the effect is akin to a change of 0.14 standard deviations. In SI Appendix 7, we report the same models using Jaccard similarity—an alternative measure of linguistic similarity, which captures whether elites incorporate new words from the opinion reports into their speeches. As SI Table A3 shows, the overlap between political speech and the research reports increases by between 1 and 1.5 percentage points after a report is issued. Overall, we therefore interpret this as evidence that the public opinion reports, indeed, affect political speech in a substantively meaningful way.

Does the increase in cosine similarity capture agenda setting? While caution is warranted, four reasons undergird this interpretation. First, the analysis focuses exclusively on substantively meaningful words (stop words are excluded). This ensures that the increase in linguistic similarity is of substantive relevance. Second, to demonstrate this more rigorously, in SI Appendix 6 we show that the increase in similarity is driven by substantively meaningful words (following a method proposed by Egesdal, Gill, and Rotemberg (n.d.)). For instance, we show that speeches on social policy more frequently rely on the word *social* and refrain from using the word *law* as a result of the dissemination. Third, we do not observe outright plagiarism. Indeed, as SI Appendix 5 shows, politicians are careful not to quote verbatim from the reports. The reports are thus not used as a mere rhetorical "stockpile." Fourth, the effect sizes are substantively meaningful. If politicians merely used the reports to change their wording, one would arguably not expect such noticeable increases in similarity were it not for a change in the topics they address.

That said, we must reiterate that the analysis matches speeches and reports that address the same broad political topics. We do so because a window of a few days is too small to realistically affect cabinet members' broader speech agenda (we confirm this in Table 3, which we discuss below). As much was relayed to us in qualitative interviews. The chancellor's schedule is set months in advance. A speech scheduled to be delivered in parliament on, say, economic policy cannot be changed to a speech on culture—at least not at such short notice. The public opinion reports can thus only plausibly affect the intensive margin of politicians' speech agendas, not the extensive margin. Agenda setting thus seemingly takes a more nuanced form: Politicians adjust their speeches but within reason. This means, for instance, that a scheduled speech on economic policy, after being exposed to public opinion, changes its focus from growth to taxation. Alternatively, a speech on foreign policy may switch its focus from aid to migration.

TABLE 2 Effects on Cosine Similarity

	Cosine Similarity	
	(1)	(2)
Exposure	0.0137**	0.0128**
	(0.0066)	(0.0057)
Covariates	No	Yes
Observations	5,684	5,684
Mean of DV	0.1263	
SD of DV	0.0976	
Effect size in SD	0.1413	0.1319

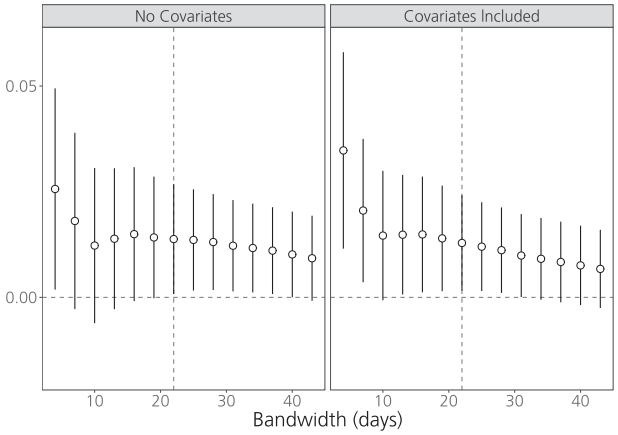
Note: The table reports results from a local linear regression around the release of the opinion reports (optimal bandwidth of 22 days; Equation 1). The outcome is the cosine similarity between reports and speeches. The sample is limited to pairs where both speech document and opinion report address the same topic. In Model 2, all covariates reported in Table 1 are included. Standard errors in parentheses are clustered by speech document and by opinion report.
*p<.1; **p<.05; ***p<.01.

Robustness

Before exploring whether exposure to public opinion also affects politicians' substantive positions, we present five robustness tests.

Bandwidth Sensitivity. In a first step, we let the RD bandwidth vary. Figure 3 provides clear evidence that the dissemination of reports leads to an increase in linguistic similarity no matter which bandwidth is chosen. Even the smallest computationally feasible bandwidth (4 days) shows statistically significant effects. Again, effect sizes

FIGURE 3 Effects on Cosine Similarity (Bandwidth Sensitivity)



Note: The figure plots coefficients and standard errors of RD regression of the cosine similarity outcome on time distance (Equation 1). The x-axis indicates the bandwidth used for the RD estimation. The y-axis plots the estimated effect size and the corresponding standard errors, clustered by speech and report. The vertical dashed line represents the optimal bandwidth (22 days).

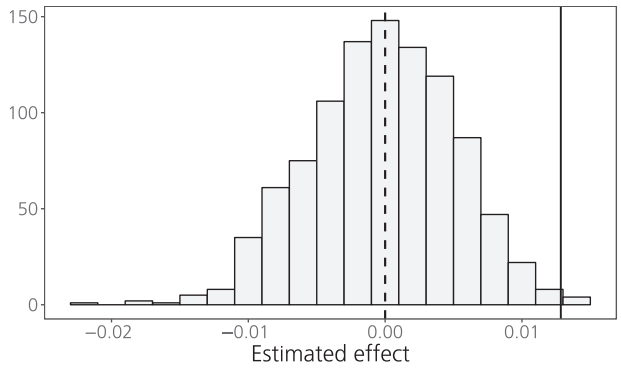
are similar when including covariates (right-hand figure). The figure also shows that the effect slowly fades out as we increase the bandwidth. This likely demonstrates that the effect of the reports is relatively short-lived, which one would expect given the fast-paced nature of politics.

Randomization Inference. In a second step, we implement a version of Fisher's exact test by rerandomizing the "treatment" (i.e., the dissemination date of the public opinion report). To do so, we first generate a list of possible report release dates. Thereafter, we sample new dates from this list with replacement for each report and calculate updated values for the time distance between the speech documents and the reports. We then re-run the benchmark RD model with the updated time distance, including all covariates, and store the corresponding LATE estimate. We repeat this procedure 1,000 times using the optimal bandwidth of 22 days. In Figure 4, we plot the distributions of the test statistic using the rerandomization procedure. For the optimal bandwidth, we find that only 1.4% of all iterations see a treatment effect that is greater in absolute magnitude than what we see in the actual data (i.e., $P(|\tau^{RD}| > \hat{\tau}^{RD}) = 0.014$).

Placebo 1: Nonmatching Topics. In a third step, we construct a placebo test that relies on the intuition that one should *not* observe treatment effects when the public opinion reports and the speech document address different topics. Given our RD design with a window of a few days, such an extensive margin effect on speech agendas would be implausible. As we argued, cabinet members' speeches are planned months in advance. Thus, if one were to observe treatment effects, this would likely imply that unobserved factors such as general shifts in speech—not exposure to the public opinion reports—are responsible for the observed changes. To construct such a test, we use the fact that we know a given opinion report's as well as a given speech document's topic. We then reestimate the RD model, matching opinion reports and speech documents that do *not* address the same broad topic. For example, this means matching a research report on education to a speech on foreign policy. The results from the placebo test are provided in Table 3. As can be seen, the estimated coefficient is virtually zero. Similarity between politicians' speech and public opinion reports does *not* increase after the dissemination of reports when the two documents address different topics.

Alternative Speech Measure. In a fourth step, we assess whether exposure to public opinion also affects an alternative measure for political speech: speeches delivered in the German parliament. Specifically, we use data provided by Rauh (2015), which cover all speeches given by members of the German government from 2009 to 2013. We focus on parliamentarians who are part of the executive and were thus plausibly exposed to the reports (for details, see SI Appendix 8). We use these data to reestimate our benchmark regression. As Table 4 shows, we see a similar treatment effect for this alternative speech outcome. In-

FIGURE 4 Randomization Inference (Cosine Similarity)



Note: The figure plots the distribution of the LATE from the benchmark RD model when rerandomizing report release dates 1,000 times. The dashed vertical line indicates $\hat{\tau}^{RD} = 0$. The solid vertical line indicates the observed LATE using the original data (see Table 2). The corresponding p-value is .014.

TABLE 3 Effects on Cosine Similarity for Unmatched Topics (Placebo)

	Cosine Similarity	
	(1)	(2)
Exposure	0.0003	0.0015
	(0.0031)	(0.0027)
Covariates	No	Yes
Observations	27,233	27,000
Mean of DV	0.0991	
SD of DV	0.0839	
Effect size in SD	0.0036	0.0179

Note: The specification follows Table 2; the sample is limited to pairs where speech document and opinion report do *not* address the same topic.

terestingly, the estimate is twice as large as the benchmark regression estimate (Table 2). This may be due to the fact that parliamentary speeches receive significant attention in the media, which makes it particularly worthwhile for elected officials to adjust their speech agenda.

Placebo 2: Opposition Members. In a final step, we use the parliamentary speech data to construct an additional placebo test. Specifically, we assess whether the dissemination of opinion reports affects speeches given by opposition party members. To create a valid counterfactual, we focus on members of parliament (MPs) who would plausibly be part of the cabinet—and thus be exposed to the reports—were their party in power (see SI Appendix 8 for details). Crucially, since opposition party members are not exposed to the classified reports, we do *not* expect their speech to change. If, however, there are unobserved shocks, one would expect opposition party members to change their speech like members of the executive. The results from the RD estimation are presented in Table 5. As can be seen, there is no significant treatment effect. Opposition party members do not change their rhetoric markedly after reports are issued.

TABLE 4 Effects on Cosine Similarity for Parliamentary Speeches

	Cosine Similarity	
	(1)	(2)
Exposure	0.0469***	0.0355***
	(0.0165)	(0.0131)
Covariates	No	Yes
Observations	3,451	3,451
Mean of DV	0.0882	
SD of DV	0.1017	
Effect size in SD	0.4608	0.3492

Note: The specification follows Table 2; the sample are parliamentary speeches.

TABLE 5 Effects on Cosine Similarity for Opposition Speeches (Placebo 2)

	Cosine Similarity	
	(1)	(2)
Exposure	0.0076	0.0189
	(0.0174)	(0.0132)
Covariates	No	Yes
Observations	1,084	1,084
Mean of DV	0.119	
SD of DV	0.1185	
Effect size in SD	0.0643	0.1592

Note: The specification follows Table 4; the sample is limited to pairs where speech document and opinion report do *not* address the same topic.

Responsiveness

We have provided evidence that exposure to public opinion leads politicians to change their speeches such that they become more linguistically similar with public opinion. But does this finding merely capture a change in the topics elites adress? Or do politicians also adjust their substantive positions? To answer this question, we next turn to the hand-coded set of 2,000 report–speech pairs.

Table 6 reports the benchmark RD model using the substantive agreement outcome. Recall that trained research assistants assigned each pair a score ranging from –3 (the politician strongly disagrees with public opinion displayed in the report) to 3 (the politician strongly agrees with public opinion). The table shows that report dissemination is associated with a significant increase in substantive agreement. Model 1 demonstrates that exposure to the reports increases substantive agreement between the reports and politicians' speeches by 0.2 (S.E. = 0.08) on the 7-point scale. A precise example of substantive repositioning is Merkel's reaction to a government report, which states: "Germans are particularly worried about rising energy prices." Six days later, a press release by the government states: "She [Merkel] sympathizes with citizens' concerns about rising electricity prices."

Again, the finding is robust to a variety of sensitivity analyses. First, Model 2 in Table 6 shows that the estimate is highly similar when including all available covariates. Second, SI Figure A.15 demonstrates that the increase is

TABLE 6 Effects on Substantive Agreement

	Substantive Agreement	
	(1)	(2)
Exposure	0.1952**	0.2048***
	(0.0833)	(0.0710)
Covariates	No	Yes
Observations	214	214
Mean of DV	0.0814	
SD of DV	0.3591	
Effect size in SD	0.5436	0.5703

Note: The table reports results from a local linear regression (Equation 1) around the release of the opinion reports (optimal bandwidth of 17 days). The outcome is substantive agreement between reports and speeches. The sample is restricted to pairs where both speech document and opinion report address the same topic. In Model 2, all covariates reported in Table 1 are included. Standard errors in parentheses are clustered by speech document and by opinion report.
*p<.1; **p<.05; ***p<.01.

detectable for a wide range of RD bandwidths. Third, SI Table A5 shows that the finding is not detectable when pairing speeches and reports that do *not* share the same topic.[3] Taken together, the evidence implies that exposure to public opinion also affects politicians' substantive positions. The link from public opinion to political speech thus seemingly operates through (at least) two channels: agenda setting and substantive responsiveness.

Agenda Setting vs. Substantive Repositioning

Can we tell agenda setting and substantive repositioning apart? One concern regarding the analyses presented thus far is that the measures of responsiveness and agenda setting overlap. Specifically, one may object that if elites adjust their substantive positions, this also manifest itself in increased linguistic similarity. Although this measurement concern does not compromise the causal inferences we draw, it does highlight the difficulty of adjudicating between agenda setting and substantive repositioning. To parse the two strategies apart, we provide four pieces of evidence.

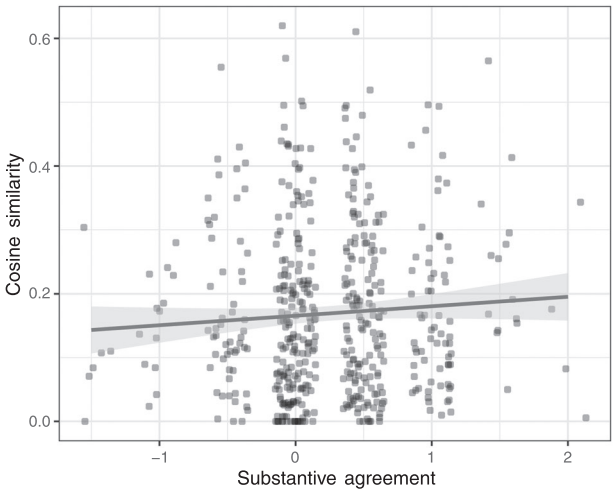
[3]Note that in order to afford this test, we hand-coded an additional random sample of 200 speech–report pairs that do *not* share the same topic. Unfortunately, we are not in a position to repeat the randomization inference robustness test. Doing so would require that we code all 2,301,431 pairs in order to ensure sufficient power around the cutoffs, which is not feasible.

First, if substantive realignment goes hand in hand with increased cosine similarity (our measure for agenda setting), one would expect the two measures to be highly correlated. If, by contrast, the substantive agreement measure is orthogonal to cosine similarity, this builds trust that the measures capture distinct dimensions of political speech. Reassuringly, Figure 5 shows no noticeable correlation between the two measures. Increases in substantive agreement (x-axis) are not accompanied by noticeable changes in cosine similarity (y-axis). Put differently, substantive repositioning does not go hand in hand with agenda setting. By the same token, a shift in agendas (cosine similarity) does not imply that politicians also change their substantive positions.

Second, one would expect significant variation across topic salience for the two strategies. Salient topics (e.g., migration) tend to polarize the electorate. Here, substantive repositioning may come across as opportunistic and create reputational costs. Politicians may thus be better off putting such topics on their agenda, rather than changing their positions. Doing so allows officials to signal that they take core issues seriously without coming across as a flip-flopper. The case is arguably different for nonsalient topics (e.g., culture). Here, agenda setting carries little weight, particularly along the intensive margin. But substantive repositioning may well win over voters without putting politicians' reputation at risk.

To assess these theoretical considerations, we explore treatment effect heterogeneity by topic salience

FIGURE 5 Correlation between Cosine Similarity and Substantive Agreement



Note: The figure shows the joint distribution for cosine similarity and substantive agreement, including the line of best fit (OLS). To ease visual interpretation, observations are jittered.

Table 7 Effects on Cosine Similarity and Substantive Agreement (Heterogeneity by Topic Salience)

	Salient Topics		Nonsalient Topics	
	Agenda Setting	**Substantive Agreement**	**Agenda Setting**	**Substantive Agreement**
Exposure	0.0145*	0.1345	0.0126	0.1944***
	(0.0075)	(0.1244)	(0.0087)	(0.0608)
Covariates	Yes	Yes	Yes	Yes
Observations	2,729	85	2,955	127
Mean of DV	0.1261	0.1199	0.1223	0.0744
SD of DV	0.0973	0.3829	0.0998	0.2797
Effect size in SD	0.1489	0.3513	0.1259	0.6952

Note: The specification follows Table 2 and Table 6, respectively; the sample is split into salient and non-salient topics.

for the two measures (the salience measure is discussed in SI Appendix 3). Table 7 supports our theoretical expectations. For salient topics, we see no substantive repositioning, but increases in agenda setting (cosine similarity). By contrast, among nonsalient topics, elites change their substantive positions but do not engage in noticeable agenda setting. While the differences between these estimates are noisy, they do buttress our intuition that the measures for agenda setting and substantive agreement are distinct and differ in theoretically plausible ways.

Third, if public opinion research leads politicians to adjust their agendas because of a "cognitive salience" mechanism (as hypothesized above), treatment effects should arguably increase when elites are repeatedly exposed to the same topic. To explore this conjecture, we proxy for repeated treatments by calculating the dissemination time gap between two reports on the same topic. The median time between reports is 33 days. In Table 8, we reestimate the benchmark RD regression, splitting the sample along the median. The results show that the exposure effect is more pronounced when another report on the same issue was released relatively recently. However, the effect is only detectable for parliamentary speeches. The finding thus provides tentative evidence that agenda setting is, indeed, more likely when politicians are exposed to multiple reports within a relatively short time frame.

Last, when politicians engage in agenda setting, one would arguably *not* expect stronger treatment effects when reports contain particularly novel information. After all, we hypothesized that the public opinion research reports lead to agenda setting by (a) communicating that the topic is relevant to the electorate and (b) making the topic cognitively salient to the politician. The novelty of the report does not factor into the equation. To proxy the degree to which reports contain new information, we calculate the cosine similarity between a report and the most recent previous report on the same topic. If a report is relatively dissimilar compared to the last report, it likely contains a greater amount of new information. We then reestimate the benchmark model separately for two subsets, defined by whether the prior report distance measure is above or below the sample median. Reassuringly,

Table 8 Effects on Cosine Similarity (Heterogeneity by Treatment Intensity)

	Government Releases		Parliamentary Speeches	
	Repeat Treat	**Nonrepeat Treat**	**Repeat Treat**	**NonRepeat Treat**
Exposure	0.0114	0.0123	0.0548***	0.0233
	(0.0077)	(0.0102)	(0.0171)	(0.0235)
Covariates	Yes	Yes	Yes	Yes
Observations	1,826	2,910	1,281	1,261
Mean of DV	0.1153	0.1303	0.0866	0.0923
SD of DV	0.0920	0.1002	0.0966	0.1081
Effect size in SD	0.1243	0.1225	0.5672	0.2156

Note: The specification follows Table 2 and Table 4, respectively; the sample is split into repeat and nonrepeat treatments, where the former comprises all cases in which two reports on the same topic were published within 33 days (median split), and vice versa.

Table 9 Effects on Cosine Similarity (Heterogeneity by Novelty of Report)

	Government Releases		Parliamentary Speeches	
	Not Novel	Novel	Not novel	Novel
Exposure	0.0118	0.0093	0.0628***	0.0435**
	(0.0106)	(0.0084)	(0.0239)	(0.0198)
Covariates	Yes	Yes	Yes	Yes
Observations	1,808	2,208	1,030	1,512
Mean of DV	0.1455	0.1044	0.1027	0.0785
SD of DV	0.1025	0.0873	0.1036	0.1007
Effect size in SD	0.1149	0.1065	0.6060	0.4317

Note: The specification follows Table 2 and Table 4, respectively; the sample is split into novel and non-novel reports based on the cosine similarity to the most recent previous report.

Table 9 shows no evidence that report novelty mediates the magnitude of the treatment effect. Put differently, the estimates across the "novel" and "not novel" subsets are not different. We must caution, however, that this analysis is not conclusive evidence that politicians entirely disregard the novelty of the public opinion reports when setting their agendas.

Generalizability

Can we characterize whether the findings generalize to other contexts? The German electoral system is unique in that it elects MPs both on the basis of plurality vote within electoral districts (majoritarian) as well as on the basis of party lists (proportional representation). Do we observe different treatment effects in the parliamentary speeches for MPs elected via party lists as opposed to MPs elected by majority vote? To answer this question, we coded whether MPs were directly elected or not using the aforementioned sample of MPs (see SI Appendix 8).

Table 10 Effects on Cosine Similarity (Heterogeneity by Electoral System)

	Cosine Similarity	
	Elected in District	Elected through Party List
Exposure	0.0291	0.0401**
	(0.0183)	(0.0177)
Covariates	Yes	Yes
Observations	1,437	1,950

Note: The specification follows Table 4; the sample is split according to whether the speaker was elected directly in his or her district or through the state party list.

We then repeat the analysis shown in Table 4, splitting the sample into directly elected MPs and MPs who entered parliament via their party's list. Table 10 shows that agenda setting is more pronounced among party list candidates. The finding is compatible with our theoretical argument insofar as the opinion reports contain information on the national median voter, which is arguably more relevant to list candidates. Plurality candidates, by contrast, face constituents whose opinion may differ substantially from the median voter. Therefore, district candidates may be less inclined to incorporate public opinion into their speeches. If taken at face value, the evidence may thus imply that agenda setting is most likely to arise in PR systems.

Still, the fact that governments around the world engage in sophisticated public opinion research makes it unlikely that the German case is an outlier (Shapiro and Jacobs 2001). What is more, the observed treatment effects permeate throughout the German political system across different parties. Angela Merkel's government is also by no means exceptional in its use of public opinion research. Chancellor Helmut Kohl, for instance, was advised by the Allensbach Institute, whereas Chancellor Gerhard Schröder had close connections with the Forsa Institute and its then CEO Manfred Güllner. This makes it less likely that the findings are the product of a highly specific case.

Discussion

This article has provided novel text-analytic evidence to assess whether public opinion affects political speech. Drawing on evidence from Germany, we found that politicians change their speech markedly when exposed to public opinion research. Not only does their speech

become more similar to the language used in the public opinion reports—a finding that points toward agenda setting—but they also adjust their substantive positions to the public's preferences expressed in the reports. The evidence thus brings clarity to one mechanism through which politicians connect with voters: speech.

Before reflecting on the substantive implications of the findings, two words of caution are in order. First, this article assessed agenda setting using cosine similarity, whereas substantive responsiveness was assessed using human coding. The latter measure, given that it was coded by humans, is less controversial. Still, reducing nuanced public opinion reports and speeches—both of which seldom touch on just one topic—to a one-dimensional agreement-scale is not without problems. Politicians may, for example, agree with some points made in an opinion report, while disagreeing with others, which likely creates measurement error. Regarding the former a critic might object that cosine similarity does not reflect a true change in agenda, but merely small rhetorical adjustments. We have tried to address this concern by showing that (i) similarity increases are driven by substantively meaningful words and (ii) speechwriters do not plagiarize from the reports. Still, we must again caution that we merely detect agenda setting along the intensive margin—a result of the local RD design.

Second, our attempt to make a causal argument deserves critical scrutiny. The consistent finding that cosine similarity and substantive agreement increase right after reports are given to cabinet members makes a causal interpretation intuitive. Yet, a skeptic might say that the observed changes are the product of a general shift in rhetoric. Although we do not believe that a few days should bring about such changes (indeed, the placebo and permutation tests paint a different picture), the criticism showcases the need to look at rhetoric in a more dynamic setting. Future research could help model such changes with greater clarity, perhaps by benchmarking political speech to speech in the media. Relatedly, a skeptic might also quibble that the finding is tautological (i.e., politicians writing the survey questions and timing the dissemination of the research). We believe that the qualitative and quantitative evidence rule out this possibility. The pronounced coefficients do underline that elected officials react to public opinion research. At a minimum, our study thus provides descriptive evidence that cabinet members systematically conduct public opinion research and subsequently change their speech.

Having discussed these caveats, we want to briefly reflect on how our research may be expanded. If German politicians, indeed, adjust their speeches to public opinion, this bears important insights for the study of representative democracy. In times of increasing polarization, a potential follow-up question is whether citizens perceive such adjustments as deceptive or manipulative. The beginning of the 2000s saw U.S. pundits lament that American politicians were more interested in responding to public opinion than in crafting their own agenda. Such "finger-in-the-wind" responsiveness was portrayed as a symbol for elected officials' lack of courage (Medvic and Dulio 2004). In the German context, Angela Merkel has been described in *The New Yorker* as "the quiet German"—a politician who silently panders to public opinion (Packer 2014). Similarly, leaked cables show that U.S. diplomats labeled the chancellor "Teflon-Merkel." Merkel's rhetoric, so the story goes, allows her to sidestep political controversy (Waterfield 2010). Sophisticated public opinion data are a double-edged sword. On the one hand, relying too heavily on public opinion research can turn political speech into a science that sidesteps truthful dialogue. On the other hand, knowledge about public preferences spurs responsiveness. These considerations showcase the need to further explore the relationship between elite speech and public opinion, mapping more fully the ways in which government officials use and interpret public preferences.

References

Becker, Sven, and Frank Hornig. 2014. "Kanzler: Regieren nach Zahlen." *Spiegel Online* 37. https://www.spiegel.de/spiegel/print/d-129095167.html.

Binzer Hobolt, Sara, and Robert Klemmensen. 2008. "Government Responsiveness and Political Competition in Comparative Perspective." *Comparative Political Studies* 41(3): 309–37.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82(6): 2295–2326.

Canes-Wrone, Brandice. 2006. *Who Leads Whom? Presidents, Policy, and the Public*. Chicago: University of Chicago Press.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma. 2019. "Simple Local Polynomial Density Estimators." *Journal of the American Statistical Association* 1–7.

Chong, Dennis, and James N. Druckman. 2007. "A Theory of Framing and Opinion Formation in Competitive Elite Environments." *Journal of Communication* 57(1): 99–118.

Cohen, Jeffrey E. 1999. *Presidential Responsiveness and Public Policy-Making: The Public and the Policies That Presidents Choose*. Ann Arbor: University of Michigan Press.

Egesdal, Michael, Michael Gill, and Martin Rotemberg. n.d. "How Federal Reserve Discussions Respond to Increased Transparency." Working Paper. https://scholar.harvard.edu/files/mrotemberg/files/fed_text_oct2015.pdf.

Eggers, Andrew C., and Arthur Spirling. 2014. "Ministerial Responsiveness in Westminster Systems: Institutional Choices and House of Commons Debate, 1832–1915." *American Journal of Political Science* 58(4): 873–87.

Franco, Annie, Justin Grimmer, and Chloe Lim. 2017. "The Limited Effect of Presidential Appeals." Working Paper. https://www.dropbox.com/s/5ym94ggasbff260/public.pdf?dl=0.

Geer, John G. 1996. *From Tea Leaves to Opinion Polls.* New York: Columbia University Press.

Gill, Michael, and Arthur Spirling. 2015. "Estimating the Severity of the WikiLeaks U.S. Diplomatic Cables Disclosure." *Political Analysis* 23(2): 299–305.

Grimmer, Justin, and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21(3): 267–97.

Hall, Wynton C. 2002. "'Reflections of Yesterday' George H.W. Bush's Instrumental Use of Public Opinion Research in Presidential Discourse." *Presidential Studies Quarterly* 32(3): 531–47.

Imbens, Guido W., and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142(2): 615–35.

Jacobs, Lawrence R., and Robert Y. Shapiro. 2000. *Politicians Don't Pander: Political Manipulation and the Loss of Democratic Responsiveness.* Chicago: University of Chicago Press.

John, Peter, and Will Jennings. 2010. "Punctuations and Turning Points in British Politics: The Policy Agenda of the Queen's Speech, 1940–2005." *British Journal of Political Science* 40(3): 561–86.

Koch, Jeffrey W. 1998. "Political Rhetoric and Political Persuasion: The Changing Structure of Citizens' Preferences on Health Insurance during Policy Debate." *Public Opinion Quarterly* 62(2): 209–29.

Lowe, Will, Kenneth Benoit, Slava Mikhaylov, and Michael Laver. 2011. "Scaling Policy Preferences from Coded Political Texts." *Legislative Studies Quarterly* 36(1): 123–55.

Lucas, Christopher, Richard A. Nielsen, Margaret E. Roberts, Brandon M. Stewart, Alex Storer, and Dustin Tingley. 2015. "Computer-Assisted Text Analysis for Comparative Politics." *Political Analysis* 23(2): 254–77.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142(2): 698–714.

Medvic, Stephen K., and David A. Dulio. 2004. "The Permanent Campaign in the White House: Evidence from the Clinton Administration." *White House Studies* 4(3): 301–18.

Bjerre Mortensen, Peter, Christoffer Green-Pedersen, Gerard Breeman, Laura Chaqués-Bonafont, Will Jennings, Peter John, Anna M. Palau, and Arco Timmermans. 2011. "Comparing Government Agendas: Executive Speeches in the Netherlands, United Kingdom, and Denmark." *Comparative Political Studies* 44(8): 973–1000.

Packer, George. 2014. "The Quiet German." http://www.newyorker.com/magazine/2014/12/01/quiet-german.

Rauh, Christian. 2015. "Communicating Supranational Governance? The Salience of EU Affairs in the German Bundestag, 1991–2013." *European Union Politics* 16(1): 116–38.

Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, and David G. Rand. 2014. "Structural Topic Models for Open-Ended Survey Responses." *American Journal of Political Science* 58(4): 1064–82.

Rottinghaus, Brandon. 2006. "Rethinking Presidential Responsiveness: The Public Presidency and Rhetorical Congruency, 1953–2001." *Journal of Politics* 68(3): 720–32.

Rottinghaus, Brandon. 2008. "Presidential Leadership on Foreign Policy, Opinion Polling, and the Possible Limits of 'Crafted Talk'." *Political Communication* 25(2): 138–57.

Shapiro, Robert Y., and Lawrence R. Jacobs. 2001. "Source Material: Presidents and Polling: Politicians, Pandering, and the Study of Democratic Responsiveness." *Presidential Studies Quarterly* 31(1): 150–67.

Slapin, Jonathan B., and Sven-Oliver Proksch. 2008. "A Scaling Model for Estimating Time-Series Party Positions from Texts." *American Journal of Political Science* 52(3): 705–22.

Vaughn, Justin S., and José D. Villalobos. 2006. "Conceptualizing and Measuring White House Staff Influence on Presidential Rhetoric." *Presidential Studies Quarterly* 36(4): 681–88.

Waterfield, Bruno. 2010. "WikiLeaks: Angela Merkel Only Leader Man Enough to 'Rule' Europe." http://www.telegraph.co.uk/news/worldnews/europe/8168873/WikiLeaks-Angela-Merkel-only-leader-man-enough-to-rule-Europe.html.

Weaver, David. 1991. "Issue Salience and Public Opinion: Are There Consequences of Agenda-Setting?" *International Journal of Public Opinion Research* 3(1): 53–68.

Wlezien, Christopher. 2004. "Patterns of representation: Dynamics of Public Preferences and Policy." *Journal of Politics* 66(1): 1–24.

Wood, Dan B. 2009. *The Myth of Presidential Representation.* Cambridge: Cambridge University Press.

Wood, B. Dan, and Han Soo Lee. 2009. "Explaining the President's Issue Based Liberalism: Pandering, Partisanship, or Pragmatism." *Journal of Politics* 71(4): 1577–92.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix 1: Topic categorization
Appendix 2: Classification of untagged speech documents
Appendix 3: Topic salience
Appendix 4: Substantive agreement
Appendix 5: Verbatim quotations
Appendix 6: Word contribution
Appendix 7: Jaccard similarity
Appendix 8: Parliamentary speech data
Appendix 9: Validating government public opinion research